DOCUMENT RESUME

ED 450 354 CS 014 290

AUTHOR Allington, Richard L.
TITLE Effects of Reading Policy on Classroom Instruction and Student Achievement. CELA Research Report.
INSTITUTION National Research Center on English Learning and Achievement, Albany, NY.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO CELA-RR-13011
PUB DATE 2000-00-00
NOTE 37p.; Paper produced for "Curriculum Wars: Alternative Approaches to Reading and Mathematics," a conference at the Kennedy School of Government, Harvard University, sponsored by the Program on Educational Policy and Governance (October 1999).
CONTRACT R305A60005
AVAILABLE FROM National Research Center on English Learning and Achievement, University at Albany, State University of New York, 1400 Washington Avenue, Albany, NY 12222. Web site: http://cela.albany.edu/policy/index.html.
PUB TYPE Information Analyses (070)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *Educational Policy; *Educational Practices; *Educational Research; Elementary Education; *Politics of Education; *Reading Instruction; *Student Evaluation
IDENTIFIERS *Policy Implementation; Policy Implications; Policy Issues

ABSTRACT

Every 30 years or so, a very public yet personal debate about the nature of appropriate reading instruction emerges in the media and in the policy talk in legislative venues. It manifests itself in a "code-emphasis" vs. "meaning-emphasis" dichotomous debate that wears on the teaching profession at the same time that it sells newspapers and advances political careers and agendas. At the present time, the country is enmeshed in the latest incarnation of the longstanding debate. This report contends that the recent research on educational policymaking and, especially, the research on policy implementation and impacts provides at least a glimpse of the likely impact the flurry of recent policies may have on educational practice. And that glimpse suggests that current educational practice will survive largely unaltered except at the margins. The report first examines the recent policymaking context, the new national emphasis on content standards. It then discusses in depth various studies exploring the effects of policymaking in beginning reading instruction. The paper finds that most studies of the effects of educational policymaking have investigated only the first aspect of the policy logic: how policies impact instruction. Most studies of policy implementation never actually address the second aspect of policy logic: that the changed instruction will produce enhanced student achievement. The paper then suggests how the effects of policymaking on student achievement might be studied, and it discusses who gets assessed when estimating effects of policy on student achievement. Contains 2 notes and 72 references. (NKA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 450 354

Effects of Reading Policy on Classroom Instruction and Student Achievement

Richard L. Allington

CELA Research Report Number 13011

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

CS 014 296

CELA RESEARCH REPORT

EFFECTS OF READING POLICY ON CLASSROOM INSTRUCTION AND STUDENT ACHIEVEMENT

RICHARD L. ALLINGTON

CELA RESEARCH REPORT NUMBER 13011

NATIONAL RESEARCH CENTER ON

ENGLISH LEARNING & ACHIEVEMENT

THE UNIVERSITY AT ALBANY ■ THE UNIVERSITY OF WISCONSIN-MADISON

THE UNIVERSITY OF GEORGIA ■ THE UNIVERSITY OF WASHINGTON

Effects of Reading Policy on Classroom Instruction and Student Achievement

Richard L. Allington
National Research Center on English Learning & Achievement
University of Florida

National Research Center on English Learning & Achievement
University at Albany
State University of New York
1400 Washington Avenue, Albany, NY 12222

Report Series 13011
http://cela.albany.edu/policy/index.html
2000

Paper produced for *Curriculum wars: Alternative approaches to reading and mathematics*, a conference at the Kennedy School of Government, Harvard University, sponsored by the Program on Educational Policy and Governance, October 1999.

National Research Center on English Learning & Achievement
University at Albany, School of Education, B-9
1400 Washington Avenue, Albany, NY 12222
http://cela.albany.edu/
518-442-5026

The Center on English Learning & Achievement (CELA) is a national research and development center located at the University at Albany, State University of New York, in collaboration with the University of Wisconsin-Madison. Additional research is conducted at the Universities of Georgia and Washington.

The Center, established in 1987, initially focused on the teaching and learning of literature. In March 1996, the Center expanded its focus to include the teaching and learning of English, both as a subject in its own right and as it is learned in other content areas. CELA's work is sponsored by the Office of Educational Research and Improvement (OERI), U.S. Department of Education, as part of the National Institute on Student Achievement, Curriculum, and Assessment.

This report is based on research supported in part under the Research and Development Centers Program (award number R305A960005) as administered by OERI. However, the contents do not necessarily represent the positions or policies of the Department of Education, OERI, or the Institute on Student Achievement. All Center reports are peer reviewed before publication.

00/11-13011

Effects of Reading Policy on Classroom Instruction and Student Achievement

Richard L. Allington

There seem few contested areas of education that can equal the "Reading Wars" in sheer longevity (Ravitch, 2000), vitriol (Boyd & Mitchell, 2000), and importance (Pearson, 1997). Every 30 years or so, it seems (Langer & Allington, 1995), a very public yet personal debate about the nature of appropriate reading instruction emerges in the media and in the policy talk in legislative venues. It manifests itself in a "code-emphasis" vs. "meaning-emphasis" dichotomous debate that wears on the teaching profession at the same time that it sells newspapers and advances political careers and agendas. It seeks to affect policy, and hence has the potential to affect instruction. At the present time, we are enmeshed in the latest incarnation of this longstanding debate. Once again we find policy makers crafting mandates for particular types of instructional materials and particular pedagogies. But to what end? The recent research on educational policy making and, especially, the research on policy implementation and impacts provides at least a glimpse of the likely impact the flurry of recent policies may have on educational practice. And that glimpse suggests that current educational practice will survive largely unaltered except at the margins.

The recent policy making context

A most significant but hardly discussed shift in American education policy is embedded in the new national emphasis on content standards. That is, that the new standards apply to all students, not just the college bound. The goal of achievement of high academic standards for all students represents, perhaps, the most significant shift in educational policy of the 20th century. For most of the century educational policy was dominated by the expectations of the "normal curve." High academic achievement was expected only from a small proportion of the student

population – perhaps 25 percent of all students. Educational practice, followed a "slow it down" policy for students whose intellectual capacity was deemed "below average" (Allington, 1991). This policy meant reduced academic expectations and a watered down core curriculum for many students. When high-stakes assessments were initiated (1970s), they were minimum competency assessments designed to ensure that most students achieved some very basic curriculum goals. But meeting even these minimum goals – a third grade reading level at the end of sixth grade, for instance – proved impossible for some students. That is, no state education agency ever reported all students achieving the minimum basic skill standards, though most states reported a slow steady progress toward that goal. Nonetheless, in most states at least one in five students failed to demonstrate even the basic skill proficiencies.

Perhaps because progress in meeting the older basic standards was proceeding ever so slowly, educational policy makers dramatically shifted the goals. The most influential move may have been the modification of the National Assessment of Educational Progress (NAEP) along with the concomitant development of the new reporting policies (Pellegrino et al., 1999). The reading assessment was reconfigured with an emphasis on assessing "thoughtful literacy," primarily through extended response items that required students to summarize, analyze, and evaluate the information provided in the longer, authentic texts that now comprised the reading test. The modification of the NAEP texts and tasks produced a new sort of assessment that was substantially different from the multiple-choice focused traditional standardized test of reading achievement. It was a shift from locating and matching information found in the test texts to tasks that focused more on student understanding of the texts and their ability to organize and evaluate the textual information they encountered.

The shift from simply reporting average student performances on the NAEP to reporting achievement in terms of the percentage of students who achieved the newly created proficiency standards produced dramatically more negative media accounts of American student achievement (Berliner & Biddle, 1995). American students' reading achievement did not actually decline, but only about a quarter of the students were reported meeting the new "proficient" standard and fewer than half met the "basic" standard. Of course, the new benchmarks were simply wishful thinking on the part of policy makers – wishful in that there was no evidence that such standards could be achieved by all students (or even most students) – and the process of standard setting was, itself, deemed substantially flawed by the National

Research Council (Pellegrino et al., 1999) and the General Accounting Office (GAO, 1993). But promoting higher achievement was high on policy makers' agendas.

A variety of proposals were developed to foster achievement of these new standards by all students. Thus, the federal report, *Prisoners of Time* (National Education Commission on Time and Learning, 1994), proposed extending the school day and school year for lower-achieving students in order to provide the additional instructional time supposed needed by some students to acquire the desired proficiencies. The report noted that, traditionally, instructional time was a constant across students and, thus, learning outcomes varied. In seeming recognition of the tradition of the normal curve, the report argued that to achieve high common academic standards would require that schools provide some students with larger quantities of instructional time through after-school and summer school programs. This represented a shift from the "slow it down" policies that had dominated educational planning (and expectations) for students who found school learning more difficult than many of their peers.

Shortly thereafter the National Commission on Teaching and America's Future (Darling-Hammond, 1997) released their report, *Doing what matters most: Investing in quality teaching*, which argued that achieving the new standards would only occur if there was a substantial investment in modifying both pre-service and inservice teacher preparation and support. At the crux of the argument was the assertion that few teachers knew how to create classroom environments that fostered the new standards and few school organizations were designed to foster such learning. A second, and implicit, theme of that report was that doing more of the same – adding a few extra hours of traditional instruction each week – was unlikely to produce the desired broader achievement of the new higher standards.

At the same time such proposals were being released there was a flurry of policy making at the state and federal levels targeted to improve reading achievement. The reauthorization of both the Improving America's Schools Act (IASA) and the Individuals with Disabilities Education Act (IDEA) stimulated legislative actions that both upped the ante for programmatic demonstrations of positive effects on student achievement and also broadened the population of students who were to achieve these standards. Testimony provided during both reauthorization debates featured an emphasis on using "scientifically proven" practices and "systemic reform" models to better deploy federal funding (e.g., Sweet, 1997). Passage of the Obey-Porter Comprehensive School Improvement Act restricted funding to those agencies targeting "proven

programs." Added into this mix was a steady push for allowing federal funds to be used in any of several plans to increase "parental choice" through the use of federal funds as vouchers or to introduce "free market" economic principles in support of the development of charter schools.

In 1997, federal legislation was passed that created a National Reading Panel charged with the responsibility to "assess the status of research-based knowledge, including the effectiveness of various approaches to teaching children to read." In 1998 the National Research Council (NRC) released its report, *Preventing reading difficulties in young children* (Snow, Burns & Griffin, 1998) which called for "truce" in the "reading wars" and offered what was dubbed a "balanced" view of the nature of effective early reading instruction.

The NRC report was followed by passage of the Reading Excellence Act (REA) that provided competitive funding for innovations developed from "scientifically based reading research." In addition to the federal actions, legislation concerning early reading instruction was also on the agenda at the state level with over 100 bills legislating particular sorts of beginning reading instruction introduced during the 1990s (Paterson, in press). Often similar themes – research-based, explicit and systematic teaching, high standards – were mirrored in this legislation. In much of this policy making one can locate the assumption that there exists a research base on effective reading instruction that has been largely ignored by curriculum developers, teacher educators, administrators, teachers, and policy makers.

However, little evidence has been provided that supports the assertion that research-based principles have been ignored in the design of beginning reading instruction or in the preparation of elementary school teachers. On the other hand, there has been a masterful public relations campaign that has encouraged a media consensus supporting such assumptions (Allington, 1999; Allington & Woodside-Jiron, 1999; Coles, 2000; Dressman, 1999).

I have attempted to sketch key policy making activity focused on improving American reading achievement simply to illustrate the intense interest in this policy arena. I would also note that much of the legislation and policy making has focused almost solely on redesigning beginning reading instruction. That is, little interest has been shown in the nature of reading instruction provided older elementary, middle school, or high school students (or adults). Personally, I find this early emphasis odd if for no other reason than the fact that American elementary students fare far better in international comparisons of students' reading achievement (and also in math and science achievement) than do older students (Elley, 1993). The fact that

American nine-year-olds ranked second in the world on reading proficiency (while older students ranked far lower) in the most recent international assessment suggests that elementary reading instruction in this country is doing relatively well compared to international peers. Likewise, there have been significant improvements in fourth grade students reading performances on the NAEP since 1990. At higher grade levels the gains are neither as consistent nor as large as those achieved by elementary school students (Donahue, Voelkl, Campbell & Mazzeo, 1999).

These data might suggest that policies on beginning reading implemented earlier were having a positive effect on achievement. This would include the shift (between 1989 and 1995) away from more heavily scripted instructional approaches with their emphasis on traditional, segmented skills teaching to beginning reading materials that emphasized an integrated approach to reading and language arts curriculum design along with an emphasis on reducing vocabulary controls so that better quality texts could be used in beginning reading instruction. Thus, between the 1989 California and the 1993 Texas reading textbook adoptions, virtually all of the large textbook publishers shifted to curriculum materials that offered integrated language arts instruction with an emphasis on providing students with access to high-quality children's literature as required by the curricular mandates in those states. By the 1994-95 school year, almost all commercial curriculum materials reflected this trend. In addition, there was a small movement away from relying primarily on commercial curriculum packages as the framework for delivering reading and language arts instruction. In a series of surveys of elementary teachers it was consistently reported that about 20 percent of teachers made little, if any, use of commercial materials, relying, instead, on children's literature as the primary curriculum (Baumann, Hoffman, Moon & Duffy-Hester, 1998; Canney, 1993; Strickland & Walmsley, 1993). About 80 percent of the teachers reported they did use commercial reading series, though most reported some blending of use of both the reading series and children's books in their lessons.

So, if American elementary students' reading achievement has been rising, and it has, why all the concern about changing beginning reading instruction on the part of policy makers? There are four explanations that seem most viable to me. First, while American elementary school students' reading achievement has been improving, there remain large numbers of children who still are failing to acquire real reading proficiency (as is the case in all nations). These children are most often poor, and because children from ethnic minority groups are more often poor,

minority children are overrepresented in the pool of lower achieving students. In addition, the largest concentrations of poor children reside in our urban centers, the major media markets. Thus, the difficulties that urban school districts, especially, have in promoting high levels of achievement have been highlighted in the media.

Second, the widespread misunderstanding of the reconfiguration of the NAEP and the shift in reporting results created the impression that reading achievement has declined. That is, when federal agencies report that approximately half of the students tested failed to achieve the "basic" level on the NAEP in reading, public concern is heightened (the difficulties with the proficiency levels notwithstanding).

Third, many of the earlier reforms violated the "grammar of schooling" (Tyack & Cuban, 1995). When Johnny stopped having weekly spelling tests, when "emergent" writing was fostered and the misspellings were not marked in red ink, when talking about the story read replaced many of the worksheets that children had long labored over, when cooperative activities began to replace competition, many parents questioned the legitimacy of the new instructional designs. The public preference for educational design would likely have beginning reading instruction that looked more traditional, more like what most adults had experienced. When the media misreported the low ranking of California on the NAEP state-by-state comparisons[1] to mean that reading achievement in that state had experienced some precipitous decline with the advent of this new instructional design, public concern about the effectiveness of the integrated, literature-based curriculum plans emerged. These concerns were particularly evident in organizations of "cultural conservatives" who had long worried about the direction of American education, and American society as well (Gaddy, Hall & Marzano, 1996; Paterson, 1998).

Fourth, and finally, there is the potential for substantial profit that would accrue to some corporations and some individual reform advocates if particular policies are put into place. The potential profits might come from any of several sources. For instance, the four corporations that now dominate reading textbook publishing (McGraw-Hill, ScottForesmanAddisonWesley, Houghton-Mifflin, Harcourt) stand to profit from increased sale if the new basal readers series move again into a dominating instructional role and schools replace existing series with the newer ones that reflect the manufactured research consensus. In addition, the reform advocates identified with curricular materials produced by those publishers would also benefit financially from increased sales revenues.

Corporations like the Washington Post Companies (Machan, 2000) stand to benefit from an increased interest on the part of middle-class parents who want private tutoring to help their children meet the new, higher literacy standards (the Post owns Kaplan, Score and other such tutoring ventures). Advantage Schools, Dreamcatchers, Sylvan, and other companies providing for-profit schooling and other educational services will profit if confidence in public education is diminished and alternatives such as charter schools, vouchers, and tuition waivers are funded with public education funds. In other words, corporate America seems to have realized the substantial profit potential that exists in the education industry (Vine, 1997).

But, in truth, no single-faceted explanation adequately explains the current interest in policy making concerning beginning reading instruction. Suffice it to say that literacy instructional policies are being promoted that are not reliably supported by the research (Allington, 1999; Coles, 2000; McQuillan, 1998; Pearson, 1997; Pressley & Allington, 1999; Taylor, 1999) and often the new policies are replacing older policies with little serious attention to the potential impacts of the policies being replaced or documenting the impacts of the new policy making. But an interesting question is whether all this policy making has the intended effect on classroom practice?

POLICY LOGIC

It seems that educational policy making proceeds on the assumption that implementing particular policies will have some intended effect. The policy logic for shaping beginning reading instruction seems to go something like this:

> If certain forms of pedagogical knowledge, instructional methods, and curriculum materials are mandated by policy, then 1) teachers will offer a particular type of beginning reading instruction that 2) will positively enhance student reading achievement in the both the short- and longer-term.

But is there evidence that would support that logic? And if not, what sort of studies might provide such evidence? What is the nature of the effects of policy on educational practice? And

what difficulties will policy researchers have in garnering evidence that would support the attainment of intended effects? It is to these questions that I now turn.

Studies Exploring the Effects of Policy Making in Beginning Reading Instruction

One fascinating aspect of policy making is that it seems that few policy makers are actually very interested in tracking the impact of the policies they create. Few pieces of educational legislation provide funds for the rigorous study of the effects of the policy being promoted. It is true that some legislation includes funding for some sort of evaluation component but rarely is this funding sufficient to examine the intended effects in any rigorous manner. The federal Title I legislation, for instance, has historically required an evaluation component, but the funding available has been miserly and the resultant evaluation studies less than particularly useful (Slavin, 1987). The same holds true for the IDEA and the REA funding. Rarely does state education legislation provide any such funding.

So bills have been passed mandating that all teachers pass a test of knowledge of phonics terms and principles (Paterson, in press). But there are no studies that suggest that such legislation will result in 1) teachers who teach differently, nor any that suggest 2) that the different sort of teaching that was imagined by the advocates of the policy produces improved reading achievement. And the legislation typically provides no funding to examine the potential relationship between policy and practice.

Likewise legislation has been passed mandating the use of particular pedagogical approaches or materials in beginning reading with little or no funding to examine whether the mandates 1) effectively altered beginning reading instruction and 2), if so, whether the altered instruction resulted in improved achievement, short- or longer-term.

Fortunately, the U. S. Department of Education's Office for Educational Research and Improvement (OERI) has been interested in such questions and has stepped in to fill the void in funding for studying the impacts of educational policy making. Over the past decade OERI has funded a number of studies on the relationship between policy and practice, and some of those studies have focused on policies related to reading instruction.

Rand Change Agent Study. Perhaps the classic study in this area is the Rand Change Agent Study (McLaughlin, 1991) that examined the implementation of federally funded education initiatives beginning in the late 1970s. This multi-year, multi-site study found that "it is exceedingly difficult to change practice" (p. 147). There was little evidence that federal initiatives were implemented consistently or faithfully. Central to the problem was the limited capacity that many school organizations had for implementing the expected changes in educational programs. In other words, limited expertise hindered local education agencies even when federal funds were allocated to support innovation in the design and delivery of intervention programs for lower-achieving children and youth. McLaughlin (1991) concluded, "Policy can't mandate what matters. What matters most to policy outcomes are local capacity and will" (p. 147). Timar and Kirp (1987) summarized the problem similarly: "Excellence cannot be coerced or mandated" (p. 309). They argued that institutional capacity for implementing new educational initiatives had been seriously overestimated (and, alternatively, the difficulty of building local capacity had been underestimated) by policy makers.

In many respects the Rand study reiterated Wildavsky's earlier (1979) admonition: "Telling people they have not achieved intended objectives does not necessarily help them discover what should be done" (p.7). He noted that if "school districts and their teachers, parents, and administrators really knew what would work and what wouldn't . . . and refused to put successful strategies into effect, they would deserve condemnation" (p. 312). Of course, the problem was that neither school district personnel nor policy makers had any clear evidence on just what would alleviate the problem of underachievement, though, as Timar and Kirp (1987) noted, policies had been crafted, seemingly from anecdotal evidence – from a "handful of success stories" (p. 314) to use their words. The problem as noted by Adam Urbanski (1991), president of the teachers' union in Rochester, NY, was that the needed "change in education is difficult. It means doing things differently, not just doing longer and harder what we already do . . . The challenge of reform is not merely to buttress the schools we have but to invent schools we've never had" (p. 29). And the available research rarely offers much evidence on what such schools might look like.

Much of the early policy implementation work was well summarized by Cohen and Spillane (1992), who argued that American traditions of decentralized educational decisions and distrust of centralized authority often undermined attempts to reform from a distance. But they also noted

that there was little evidence that the press for more standardization in educational planning would actually result in higher achievement, especially achievement of higher-order learning goals. They concluded that the evidence suggested that even with common curriculum goals and common textbook usage, teachers varied substantially on what was taught and how they taught. Ultimately, they argued, a "logic of confidence" replaced good evidence of rational relationships between educational resources, instructional processes, and achievement outcomes. In other words, there was little actual evidence of what worked to produce broad achievement of the new higher standards and less evidence that what worked in one locale was transportable to other locales.

Consortium for Policy Research in Education studies. The early studies of policy effects produced a second wave of educational policy studies. The largest effort was carried out by the Consortium for Policy Research in Education (CPRE), with one series of studies examining the efforts to reform elementary school reading instruction (Elmore, Peterson & McCarthy, 1996). Three schools that had restructured in an attempt to foster higher achievement, particularly focused on developing children's thoughtful literacy proficiency, were studied extensively. The research team conducted sustained classroom observations combined with interviews and analyses of student learning in an attempt to better understand the efforts to reform. What they found was substantial reorganization of the schedule and patterns of assignment of students to classrooms but only modest modifications to the core instructional activities. In other words, the schools had moved to multi-age assignment of students to classrooms, to larger blocks of time allocated to specific subject matter, and often to a more integrated approach to curriculum. But these changes in organizational patterns had minimal impacts on teaching. In most cases the teachers continued to teach in ways that they were familiar with – traditional teaching.

Elmore (1996) summarized the findings by noting that, "It is unlikely that teachers who are not intrinsically motivated to engage in hard, uncertain work will learn to do so in large, anonymous organizations that do not intensify personal commitments and responsibilities" (p. 25). But changing instructional practice is demanding, risky business in a high-stakes assessment environment (Johnston, Allington, Guice & Brooks, 1998; Miller, 1995). Elmore noted that changing aspects of schooling that are distant from the core activities of teaching were more likely to be readily adopted than changes in those aspects central to teaching and learning.

The Michigan Educational Policy and Practice studies. This series of studies involved analyses of how school district personnel interpreted and implemented Michigan's new reading curriculum framework and the aligned high-stakes assessment that accompanied it. A central aspect was developing a better understanding of local variation in the implementation of the same policy mandate. Spillane (1998) points to the importance of individuals' beliefs about the nature of appropriate reading instruction in their interpretation and implementation of the policy mandates. In addition, different participants had access to different information sources on the policy mandate that influenced their understanding of what was to be done. But different participants who had access to the same sources of information often interpreted the policy differently based upon the particular situation they occupied and the beliefs they held about teaching and learning to read.

Jennings (1996) suggested that it was more appropriate to consider policy implementation as "an incident of teaching and learning, rather than as a process by which ideas are filtered through an educational system and enacted by practitioners" (p. 107). In this case, the curricular framework was aligned with the new assessments but the framework itself had little to offer teachers about the selection and use of texts during reading instruction. This assessment-driven reform relied on teachers inventing instructional practices that led to improved achievement on the new assessments. And the state's attempts at providing teachers with models of this new instruction, in Jennings' view, were largely fraught with problems as policy makers themselves disagreed on what instruction might look like and the explanatory sessions that were offered typically employed "old pedagogies and pieces of changed reading practices, not a coherent picture" (p. 17).

Standerford (1997) studied implementation of the Michigan reforms in two school districts over a 3 1/2 year period. She reported that "although all participants were making instructional changes, those changes were uneven and scattered within and across classrooms and districts" (p. 84). The teachers reported, and were observed, attempting to understand the new model for reading instruction. But none felt that adequate information or support had been provided, and changes attempted were rarely inconsistent with the teacher's prior knowledge or beliefs about the best ways to teach reading.

Again, these studies point to the difficulty of faithful policy implementation. Even when curricular frameworks are aligned with assessments and when a substantial coordinated effort is made to inform the public and the profession about the new policies, faithful implementation was difficult, at best. McGill-Franzen (2000) summarizes these studies by noting that "policymakers must find ways for policies to educate. . . . teachers need opportunities to learn from policy. But what they learn from these opportunities depends also on what they bring to policy – their own knowledge and beliefs" (p. 900).

The Texas basal adoption studies Texas has long had state-level control of the textbook adoption process. In 1993 Texas required the new reading series to 1) provide opportunities to read connected text rather than isolated skill workbooks, 2) provide an anthology of quality unabridged children's literature, 3) integrate reading with writing, listening and speaking, along with 4) a systematic presentation of phonemic awareness (Hoffman, et al., 1998). Analyses of the new series selected for adoption in Texas revealed that, compared to the older, skill-based series in use, vocabulary control had been largely abandoned in the newer series; thus, the new series were found to be more engaging even as they posed greater decoding difficulties for beginning readers. In addition, new procedures for introducing the story, for pacing the lessons, for grouping students, and for assessing student progress were also evident in the newer series (Hoffman, et al., 1995). Thus, the adoption of the newer reading series posed challenges for teachers.

Hoffman and his colleagues (1998) studied the reading instructional beliefs and practices of teachers in four school districts while they were using the older skills-based series and after the adoption of the newer integrated, literature-based readers. They found significant differences in instructional practice in both phases of the study. That is, instructional practices represented a continuum according to the use and role of various types of curriculum materials. Some teachers never used the reader series in either phase of the study; some of these teachers relied on children's literature, others on isolated skills work materials. Other teachers followed the instructional manuals accompanying the series in both years while others dipped into the manual and the series only occasionally. Some supplemented the reader series with children's literature and others supplemented with additional isolated skills work.

In addition, after the adoption of the new reader series some teachers continued to use the older series while others used the newer series but ignored the instructional guidelines and offered lessons following the instructional model offered in older reader series. And others continued to use children's literature exclusively and let the new series sit on the shelf.

Teachers' epistemological orientations were "determining factors" in how they organized reading instruction and how they responded to the changed curricular guidelines. Changing the curricular materials had little impact on epistemological orientations. Interestingly, the teachers commonly reported the new materials had an impact on their students. But this impact was viewed as more motivational than achievement related. The teachers felt their children were no more skilled than in earlier years but that the children were more interested in reading and reading independently.

But Hoffman and his colleagues (1998) also noted that while school districts invested substantial funds in the purchase of new materials, none of those districts invested any significant funds in professional development. The typical teacher in their study participated in a single 1- to 3-hour publisher-sponsored workshop as preparation for implementing the new instructional framework.

Policy Analysis for California (PACE) studies. California policy making has been much discussed, but the most coherent and informative studies of California policy making and the effects of those policies have come from the research team at PACE. Carlos and Kirst (1997) provide a far-reaching descriptive analysis of California educational policy making across a decade. More focused on the political environment that fostered the policy making than many policy studies, they describe the struggles between the state superintendent of instruction, the state board of education, the legislature, and the governor for control of educational policy making. They conclude that the net effect was a reduced role for the department of education as legislative acts became more specific and state board directives more controlling. Carlos and Kirst note that, "California's recent struggles in carrying out a continuous and coherent education reform agenda can be attributed largely to a fractured governance structure and the partisan conflicts and alliances that arise with each election" (p. 19). California's reform effort has been a moving target that leaves teachers wondering just what the next "quick fix" might be.

Chrispeels (1997) argued that the earlier California policy initiatives (particularly the 1987 curriculum framework and the associated assessment development) had created a decade-long coherent policy environment that had an impact on practice. However, Freeman and Freeman (1998) offer a less optimistic view. They note that relatively few teachers participated in the capacity-building initiatives sponsored by the earlier California Literature Project – the primary effort to foster teacher understanding of the 1987 frameworks. They point out that while California school districts adopted the integrated, literature-based reading series that the framework called for, many teachers misinterpreted the "whole language" framework to mean whole class instruction and had all children reading the same selection at the same time. In addition, because so much money had been allocated for the purchase of the new readers, little funding was available to build classroom collections of children's books and less was available to upgrade the quality of school library collections – collections that ranked among the poorest in the nation (McQuillan, 1998).

In short, there is no clear evidence that suggests that the earlier California framework for reading was ever consistently implemented, although opponents have made much of assertions to the contrary. What we do know is that only a limited number of California teachers participated in any substantive professional development activity designed to enhance their capacity to teach in the manner envisioned in the 1987 Framework. There seem to be no studies that assessed, on any large scale, the impact of that policy making on instructional practice, just as there seems to be no such effort assessing the impact of the more recent California policy making that represents what seems to be a 180° turn in policy direction.

The Center for Literature Teaching and Learning (CLTL) studies. As educational policy makers moved to push the use of high-quality children's literature to the forefront of educational practice, OERI funded the CLTL to engage in research on the impacts of literature-based instruction. A five-year study of the implementation of literature-based reading instruction in four higher-poverty school districts was completed. This study (Allington, Guice, Michaelson, Baker & Li, 1996; Johnston, Allington, Guice & Brooks, 1998) was unique in that the design was a 2 X 2 crossed model where two districts selected literature-based basal reader series and two elected to use children's literature as the primary curriculum materials. In addition, two districts (one basal and one books) were identified as operating under centralized decision

making while the other two districts (one basal and one books) were identified as more decentralized in the decision making patterns, especially as concerned the selection of curriculum materials and frameworks. This design was intended to separate the relative impact of organizational factors from curricular factors in the study of educational reform.

As in the CPRE study (Elmore et al., 1996), much of the "reform" involved restructuring schedules and student assignments to classrooms and providing different curriculum materials. However, in the long run it seems that organizational patterns of decision making were more important than the nature of the curriculum materials. That is, in the districts where decision making was more centralized there was less change and reading achievement actually declined modestly over the five year study period. And one district mandated the use of a commercial reader series while the other mandated the use of children's literature. In the districts where decision making was more decentralized – where large teacher committees made decisions about curricular issues – more consistent and positive change was observed and small improvements in reading achievement followed.

Nonetheless, even in the more decentralized settings, the pace of change in core instructional activities was limited. There was an observed decline in time allocated to low-level skills work and an increase in the amount of reading and writing that students completed. But even the most changed classrooms still looked remarkably similar in most respects after five years of reform activity. At the same time, professional development opportunities targeted to enhance teacher instructional capacity were enormously limited in three of the four districts. For instance, in neither district implementing a children's literature-based curriculum was there any professional development focused on enhancing teacher awareness of children's books (even though teachers indicated that a lack of familiarity was a significant problem for them).

Kentucky Educational Reform Act (KERA) studies. There were others who studied educational reform and arrived at similar conclusions. Holland (1998) reported on the responses of educators to the ambitious Kentucky educational reform initiatives. She noted that, "Change is not just about policies, or programs, or promises. It is an intensely personal decision to try something new. And to work, change depends on a broad belief that doing something differently will make it better" (p. 26). But in this case, policy makers changed direction "faster than a

weather vane caught in a hurricane" (p. xxv). The shifting policy directives left teachers, administrators, and parents in a dither.

Guskey and Oldham (1997) provide a comprehensive review of the KERA policies and studies that suggests that there were literally no studies of the effects of reform policies on achievement. They also note the multiple individual KERA reforms were mandated with no evidence that the various components would actually work effectively when implemented as part of a systemic reform package. Indeed, the various components were often inconsistent one with another and implementation was uneven, distorted, and difficult. In the end, shifts to multi-age classes and the use of tables rather than desks were better documented than shifts in core instructional practice (Bridge, 1994).

Other smaller scale studies of policy implementation. Scharer (1992) reported on a study of teacher implementation of literature-based instruction following a district mandate. All of the teachers fell in the "skills emphasis" category on an inventory of teacher pedagogical orientation. She noted a gradual shift in instructional activities such that time allocated to isolated skills worksheets decreased as time allocated to independent, self-selected reading increased. Likewise basal reader series began to occupy less time as reading of children's literature occupied more. But teachers lack of familiarity with children's books, the limited opportunities for collegial interaction and visits to other classrooms, their limited repertoires of organizational and instructional strategies, their preoccupation with "coverage" of basals or books, and the difficulties with evaluating students for report cards all proved substantial obstacles to implementation of new models of instruction.

Pace (1992) studied teachers in four school districts who were involved in "grassroots" implementation of "whole language literacy instruction." She suggests that "the tension between individual innovators and other teachers may be the most important factor to address in accomplishing classroom reform" (p. 471). In every case she presents, teachers involved in the grassroots reform were powerfully effected by the responses of their peers. Half of the participants changed schools or grade levels or left teaching in order to avoid the conflicts involved in teaching differently. The remaining teachers curtailed their reforms and continued to struggle with their beliefs while continuing to use at least some traditional practices as an

accommodation to their peers. As Pace notes, little attention has been paid to the influence of peers in the professional work environment on reforming educational practice.

Summary of the Studies of Policy Implementation

Most studies of the effects of educational policy making have investigated only the first aspect of the policy logic: How policies impact instruction. The preponderance of the evidence available suggests that educational policy making rarely, if ever, reliably achieves the intended shifts in instructional processes. If the policy logic set out earlier reflects the intention of policy makers, then the evidence available fails to support that aspect of the logic that suggests that mandates produce intended instructional changes. It seems that policy makers inevitably underestimate the difficulty of fostering complex change in large bureaucratic entities such as schools. And changing schools – especially changing the core aspects of instruction – is, demonstrably, an enormously complex undertaking. Most studies of policy implementation never actually address the second aspect of policy logic: That the changed instruction will produce enhanced student achievement. Perhaps that is because the findings on implementation so consistently suggest that few policies are faithfully implemented.

How might we study the effects of policy making on student achievement?

Given the enormous interest in improving the academic achievement of American students, and the enormous investments of money, time, and individual effort that accompanies policy making in this arena, it would seem that more attention might be paid to the design and funding of studies that address the impact of policy making on student outcomes. However, such efforts will be complicated and, thus, expensive. Below I offer a brief consideration of a few of the design concerns that must be addressed in order to effectively research the relationship between policy making and student achievement.

Policy Consistency, and the Problem of Time Lag

In much of the research on educational policy making, as illustrated by the studies of reading policy making reviewed earlier, the problem of "policy collisions" – policy making that produces contradictory policy mandates – seems an important issue. That is, before the impact of one set of policies can be evaluated, a new set of policies that work at cross purposes is proposed. For instance, in both Texas and California, earlier policies on beginning reading instruction were replaced by new policies that largely voided the earlier policies. In Texas, however, reading achievement had risen steadily under the older policies, while in California there was no clear evidence on the effects of the earlier policy on student achievement (except, perhaps that reading achievement was not rising significantly). In neither case was there much good evidence that the earlier policies had actually been widely and faithfully implemented.

Efforts to change instruction through policy making are beset by the problem of time lag. In both Texas and California, for instance, the guidelines for new instructional materials are developed several years in advance of the textbook adoption in order to allow publishers the opportunity to create instructional materials that respond to the new frameworks. The 1987 California Framework led to the 1989 adoption of the integrated, literature-based reader series so that by the 1990-91 school year virtually all schools provided teachers with instructional materials that reflected the Framework mandates. But the professional development opportunities designed to build local capacity to respond appropriately to the instructional dimensions of the Framework had barely touched most teachers at that point. But the 1993 release of state-by-state comparisons on the 1992 NAEP produced concern that the new Framework was having negative effects on fourth grade student achievement! And by 1995, when the 1994 NAEP results were released, policies substantially altering the 1987 Framework were already in the works.

Likewise, in Texas the new integrated, literature-based instructional materials required by the 1990-91 textbook proclamation arrived in classrooms for the 1993-94 school year. But, by 1995 the Texas State Board of Education was already enmeshed in debates about the need for substantially altering the state curriculum framework and by 1996 statewide conferences on the topic were being held, this while student achievement continued to rise (Ellis, 1998).

It would seem that many of those interested in touting the impacts of educational policy making seriously underestimate the time lag between the emergence of new policies and when the implementation of these policies is initiated. Beyond that there is the problem of the time lag between initial implementation and widespread, faithful implementation of the policies.

Evaluating Fidelity of Policy Implementation

Little of the research to date suggests that educational policies are ever widely implemented as envisioned by the policy makers. Perhaps this conclusion has more to do with the instability of the policy environment than anything else. In other words, there have been few long-term studies (7-10 years) of policy implementation, in part because educational policy making continues rather hyperactively, with many policies replaced well short of any 10 year timeline. My point here is that it should not be surprising that studies of policy implementation early in the implementation process find that policy implementation is less faithful than hoped. It takes time for policies to reach schools and even longer to reach classrooms.

But if we were truly interested in the effects of policy making on student achievement we could design studies that evaluated such impacts, if any exist. The Hoffman et al. (1998) study offers one of the few good examples of such a study on the impact of state policies on literacy instruction. One key was documenting the nature of instruction before the new policy was implemented. The point is that without reliable information on the nature of the instructional environment before a new instructional policy is implemented it will be impossible to assess effects on practice. In the Hoffman et al. (1998) study, then, baseline data on instructional practice were gathered and then additional data were gathered in the initial year of implementation of the new instructional materials. That study suggested that fidelity of implementation varied widely. Unfortunately, there was no direct assessment of reading achievement (though teachers were asked about student growth compared to earlier years).

Nonetheless, the Hoffman et al. (1998) study is almost unique because it developed baseline information so that instructional changes could be documented. Many other studies have attempted to document policy effects but without such baseline information. And policy makers often seem to imagine such data exist. In California, for instance, policy makers seem to have

accepted assertions about the implementation and negative effects of the 1987 Framework even though the available evidence, scant as it was, suggested that assertions about widespread and faithful implementation were wrong. At a 1997 American Educational Research Association Annual Meeting session, key California policy makers admitted as much, but this seemed to not diminish their enthusiasm for the shifting policy agenda.[2]

Ideally, policy implementation and effects studies would begin with developing baseline data prior to the implementation of any new policy and continue through documentation of fidelity of implementation. Perhaps considering fidelity of implementation as a continuum would be a first necessary step in such research. I can imagine a multi-method research design that used surveys, observations, teaching instructional time logs, interviews, and other methods to estimate implementation fidelity. A stratified random sampling of teachers across a state, for instance, would provide useful evidence for estimating whether particular state policies produce the intended effect on instructional processes. But since policy implementation takes time, such studies would need to be designed as longitudinal efforts.

Evaluating Trends in Student Achievement

After documenting the process of policy implementation, including varied levels of implementation fidelity, we could then move on to attempting to estimate the effects of implementation on student achievement. Treating fidelity of policy implementation as a continuous variable in the research design should allow for reasonable estimates of effects on achievement across the range of policy implementation.

But how to assess student achievement? Currently there are few guidelines for making such decisions. One can find claims of achievement effects on reading, for instance, where children's achievement was estimated by their performance on a test of pseudo-word decoding (e.g., boj). But if policy makers are interested in children's development of higher-order literacy proficiencies (e.g , summarize the main argument and supporting details), then achievement estimates would necessarily involve assessing those proficiencies more directly. It seems that it was such an interest in higher-order literacy proficiencies that stimulated the redesign of the NAEP in the 1980s. Unfortunately, little of the policy research – or intervention research, for that

matter – has, to date, actually employed assessments of higher-order literacy in estimating literacy achievement. In other words, there exists a fundamental need for substantially better data on how particular instructional models or materials impact on the development of higher-order literacy skills and strategies.

For instance, such "proven programs" as Success for All, direct instruction, and the Open Court reader series, provide no evidence of the impact of these programs on the development of the sort of higher-order literacies that are central to the NAEP estimates of student reading proficiency. At best, the evidence points to slim positive effects on traditional standardized reading tests. (See Pressley & Allington, 1999; Venezky, 1998; Walberg & Greenberg, 1998 for discussions of the poor quality of the evidence supporting these programs.) The redesign of the NAEP was intended to provide the sorts of evidence long absent from traditional tests with their emphasis on lower-level skills of word recognition and information locating and matching. But no one has reliably linked particular sorts of instruction to NAEP estimates of proficiency, although several commentators have noted some interesting correlational data (e.g., Goodman, 1998; McQuillan, 1998). A few recent studies of classroom reading instruction have assessed achievement using assessments of higher-order reading proficiencies (e.g., Pressley et al., in press) and a few states have developed their own assessments that tap such skills (e.g., MD, MI, NY, WI) but many states (e.g., CA, NC, TX) have state assessments that tap primarily lower-order skills proficiencies.

A critical aspect of any study attempting to evaluate the effects of educational policy making on achievement, then, will be building or selecting a system for assessing student achievement of the new high academic standards and the higher-order literacy proficiencies. Ideally, such an assessment would be administered at the beginning and end of each school year. The problem with a single annual assessment is the difficulty of accounting for the demonstrated differential effects of summer reading loss. As Cooper Nye, Charlton, Lindsay and Greathouse (1996) have demonstrated in their meta-analysis of the available studies, children from lower-income families experience a substantial loss of reading skills over summer vacation periods, losses not found among children from more advantaged families.

Entwisle, Alexander and Olson (1997) demonstrated just how powerful this summer reading loss is when attempting to estimate school effects on achievement. They found no significant differences in the effects of different schools on student achievement even though the average

achievement in schools serving primarily students from lower-income families lagged several years behind the achievement of schools serving more advantaged students. The 2+ year difference in average achievement at the end of the elementary grades was wholly accounted for by a combination of initial achievement differences (observed when students began school) and the cumulative effects of the differential summer reading loss. They also noted that without the twice yearly assessments, differences in school achievement patterns are easily misinterpreted as located in differences in instructional effectiveness. They summarize their findings by noting that policy makers need to realize the enormous impact of summer reading loss in higher-poverty communities.

Additionally, attempting to sort the effects of extended time program participation (e.g., after-school, summer school) from the effects of regular education programs becomes complicated when participation in such programs is inconsistent across the student population being studied. Similarly, if parents of more advantaged students are more likely to provide children with outside tutoring, for instance, school and policy effects become muddied. But it would not be especially difficult to gather such data and estimate effects on achievement.

Who Gets Assessed When Estimating Effects of Policy on Student Achievement?

In designing a study of policy effects attention must be paid to the consistency and comprehensiveness of the student population participating in assessments designed to evaluate achievement. Two common problems exist in most current assessment plans, and both must be adequately addressed in order to reliably estimate achievement patterns over time.

First, there is the problem of student mobility. If a policy is implemented throughout a school system (state-wide or district-wide), then this problem may be largely ameliorated. However, this is only true if policy implementation is largely consistent throughout the system. But when a policy is implemented in only some schools – lower-achieving schools, for instance – it becomes more difficult to estimate policy effects. If, for instance, some schools have more stable student enrollments while others have more mobile populations, the comparisons of effects are muddied. One potential solution to this problem is to again use the idea of a continuum and disaggregate achievement data based on the numbers of year of schooling students receive in the schools

implementing the intended policy. For instance, one might reasonably expect different effects, positive or negative, on those students enrolled in a school for three consecutive years than on students who arrived at the school a few months before the administration of the achievement test.

Second, there is the problem of consistency in the assessment pool. It has been demonstrated how an increasing pool of students identified as pupils with disabilities produces artificially inflated improvements in achievement (Allington & McGill-Franzen, 1992b; McGill-Franzen, 1994). In both studies, this increased identification meant that more lower-achieving students were exempted from the state assessments, thus reducing the numbers of low-achieving children tested and, thereby, raising average achievement levels. Likewise, Allington and McGill-Franzen (1992a) demonstrated how increased retention in grade artificially inflated reports of improving achievement. In other words, retaining lower-achieving students removed them from the cohort assessment pool. Thus, achievement was misinterpreted as improving and the schools were mistakenly considered to be more effective instructionally. One solution to this problem is replacing grade-level assessment with cohort assessment. In other words, the assessment plan would continually assess and report the achievement of any cohort of students who began school in the same year regardless of what grade they were in over time.

Isolating the Effects of an Individual Policy

Finally, and importantly, studies of policy implementation and effects must attempt to isolate the impact of a particular policy on student achievement. Because a plethora of factors have been identified as having impacts on student achievement it will be important to design evaluation studies that attempt to account for factors known to be related to achievement. For instance, the recent flurry of policy making in California will make it more difficult to isolate policy effects. Class size reduction, new mandated instructional materials, ending social promotion, mandating summer school attendance for lower-achievers, enhancements in pre- and inservice teacher preparation, and increased per-pupil spending are all recent policy shifts in California education.

But if there is a general improvement (or decline) in the reading achievement of California students, sorting out any causal relationships between policies and achievement will be a

complicated affair. The available evidence suggests that class size reduction alone should improve early literacy development as should access to more expert instruction and increased per-pupil spending. And ending social promotion will artificially enhance reading achievement, at least initially, and the mandated summer school may improve achievement over the long run. However, as Lindblom and Woodhouse (1993) note, such trial and error policy making is often high-cost policy making. They suggest an approach that would promote limited policy implementation as studies of effects are completed. Perhaps one or two of the policies would be implemented on a trial basis at different sites, with a more gradual reallocation of funds as evaluation studies provided better estimates of the effects of the various policy components.

Summary: Potential Areas of Study

We have too little reliable information on the effects educational policy making has on student achievement. Such studies could be designed if policy makers are interested in the effects of their policy making. But the design of such studies will be complicated, and gathering reliable information on achievement patterns will be a relatively expensive undertaking. However, that expense is actually modest in comparison to the funding allocated for almost any educational reform.

SOME FINAL CONSIDERATIONS

Tyack and Cuban (1995), in their book-length review of American educational reform activity across the 20th century, argue that, "It is policy talk, we suggest, that cycles far more than practice" (p. 40). The research on the effects of policy making focused on reading instruction seems to support that conclusion. Just as the faithful implementation of "whole language" policies (with all the attendant problems of defining just what a "whole language" policy might be) seems to have been widely exaggerated in the policy talk, so too may policy makers' expectations for the implementation of "phonics" policies be overly optimistic. Lindblom and Woodhouse (1993) conclude, "When educators view policy demands as inappropriate, they are skilled in finding ways to temper or evade their effects" (p. 79).

The research suggests that implementation of educational policies entails individual teachers' translation of the policy and its intended impact on instructional processes. But teaching is a complex activity, and teaching well involves more than technical skill. Studies of teaching and teachers, especially expert teachers, point to the myriad of non-technical components of exemplary instruction (Pressley et al., in press; Ruddell, 1995; Spencer & Spencer, 1993).

Two issues deserve consideration here. First, even the technical aspects of teaching are rooted in beliefs about teaching and learning. The crux of the disagreements about teaching reading is located in differing beliefs about how best to teach – beliefs buttressed by particular interpretations of particular research. Individual teachers may rely more on personal experience than published research, but even that experience is influenced by beliefs about the nature of teaching. As a California teacher told McGill-Franzen and her colleagues (quoted in McGill-Franzen, 2000), "Teachers are independent and you cannot tell them how to think and feel. Because if you don't believe in a program . . . its not going to happen. Because when you go in that classroom and you close that door, you're going to teach what you feel and believe is right and what you feel and believe you can do" (p. 906). A second consideration is the question of institutional and individual capacity. This is, perhaps, the more technical end of things. But even the capacity for offering high-quality instruction is more than simply a technical matter. However, as Loveless (1998) notes, "If policy intends to help teachers who are incapable of effective instruction, state proclamations will not likely provide much assistance" (p. 300). The unfortunate truth of the matter is that materials don't teach. Instructional materials of different quality may have some modest impact on the nature and quality of teaching but only if and when teachers elect to 1) use those materials, and 2) use them in the ways imagined by the developers and policy makers. Throughout the policy implementation studies reviewed here we found the constant problem of fidelity of implementation – even with instructional materials that offered detailed designs for how instruction should proceed.

So how might policy making become more productive? In actuality we know rather little about how educational policies might be successfully developed or implemented – the research to date largely documents the difficulties of achieving faithful implementation. Given the substantial funding of public education and the costs of the various policy proposals it is surprising how little attention policy makers seem to have paid to the implementation process or

the research available. There is little evidence that either aspect of the logic that seems inherent in policy making is ever fulfilled.

For educational policy making to become more productive, better information on policy impacts will be needed. But there has been little interest amongst policy makers on reliably documenting the effects of their policy making. Perhaps this can be changed. But maybe there is less to educational policy making than meets the eye. Perhaps the real goal of policy makers is to have "done something," particularly something that public opinion polls support (Allington, 1999). In this case the policy making is an end unto itself. If this is so, we can expect a continued avalanche of educational policies that offer no coherent or continuing vision of how schools might improve.

ENDNOTES

1. McQuillan (1996; 1998) and others have provided several analyses of the California NAEP reports and noted that ranking near the bottom of the participating states is not the same as declining performance. The NAEP reading performances of California children has remained stable in the assessments since 1992 (Donahue et al., 1999, pp. 114-115). The analyses indicate that California has been plagued by low reading achievement since the passage of Proposition 13 in the 1970s. That legislation reduced education funding and effectively resulted in larger class sizes, lower teacher salaries, and more restricted access to libraries and books.

2. When questioned about the availability of reliable information on implementation of the 1987 Framework, Bill Honig and other panelists admitted that there were no data available on scope or fidelity of implementation.

REFERENCES

Allington, R. L. (1991). The legacy of 'slow it down and make it more concrete.' In J. Zutell & S. McCormick (Eds.), *Learner factors/teacher factors: Issues in literacy research and instruction* (40th Yearbook of the National Reading Conference, pp. 19-30). Chicago: National Reading Conference.

Allington, R. L. (1999). Crafting state educational policy: The slippery role of research and researchers. *Journal of Literacy Research, 31,* 457-482.

Allington, R. L., Guice, S., Michelson, N., Baker, K., & Li, S. (1996). Literature-based curriculum in high-poverty schools. In M. Graves, P. v. d. Broek, & B. Taylor (Eds.), *The first r: Every child's right to read,* (pp. 73-96). New York: Teachers College Press.

Allington, R. L., & McGill-Franzen, A. (1992a). Unintended effects of educational reform in New York State. *Educational Policy, 6,* 396-413.

Allington, R. L., & McGill-Franzen, A. (1992b). Does high-stakes testing improve school effectiveness? *ERS Spectrum, 10*(2), 3-12.

Allington, R. L., & Woodside-Jiron, H. (1999). The politics of literacy teaching: How 'research' shaped educational policy. *Educational Researcher, 28*(8), 4-13.

Baumann, J. F., Hoffman, J. V., Moon, J., & Duffy-Hester, A. (1998). Where are teachers' voices in the phonics/whole language debate? Results from a survey of U. S. elementary teachers. *Reading Teacher, 50,* 636-651.

Berliner, D. C., & Biddle, B. J. (1996). *The manufactured crisis: Myths, fraud, and the attack on America's public schools.* White Plains, NY: Longman.

Boyd, W. L. & Mitchell, D. E. (2000). The politics of the reading wars. In T. Loveless (Ed.), *The Curriculum Wars.* Washington, DC: Brookings Institute

Bridge, C. A. (1994). Implementing large-scale change in literacy instruction. In C. Kinzer & D. Leu (Eds.), *Multidimensional aspects of literacy research, theory, and practice.* (44th Yearbook of the National Reading Conference, pp. 257-265). Chicago: National Reading Conference.

Canney, G. (1993). Teachers' preferences for reading materials. *Reading Improvement, 30,* 238-245.

Carlos, L., & Kirst, M. (1997). *California curriculum policy in the 1990's: "We don't have to be in front to lead."* San Francisco: WestEd/PACE.

Chrispeels, J. H. (1997). Educational policy implementation in a shifting political climate: The California experience. *American Educational Research Journal, 34,* 453-481.

Cohen, D. K., & Spillane, J. P. (1992). Policy and practice: The relations between governance and instruction. In G. Grant (Ed.), *Review of Research in Education* (vol. 18, pp. 3-49). Washington, DC: American Educational Research Association.

Coles, G. (2000). *Misreading reading.* Portsmouth, NH: Heinemann.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66*, 227-268.

Darling-Hammond, L. (1997). *Doing what matters most: Investing in quality teaching.* New York: National Commission on Teaching and America's Future.

Donahue, P. L., Voelkl, K. E., Campbell, J. R., & Mazzeo, J. (1999). *NAEP 1998 reading report card for the nation and the states.* (NCES 1999-500). Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement, U. S. Department of Education.

Dressman, M. (1999). On the use and misuse of research evidence: Decoding two states' reading initiatives. *Reading Research Quarterly, 34*, 258-285.

Elley, W. B. (1992). *How in the world do students read? IEA study of reading literacy.* The Hague, Netherlands: International Association for the Evaluation of Educational Achievement.

Ellis, L. (1998). We'll eat the elephant one bite at a time: The continuing battle for control of literacy education in Texas. In K. S. Goodman (Ed.), *In defense of good teaching: What teachers need to know about the reading wars* (pp. 87-105). York, ME: Stenhouse.

Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review, 66*, 1-26.

Elmore, R. F., Peterson, P. L., & McCarthy, S. J. (1996). *Restructuring in the classroom: Teaching, learning, and school organization.* San Francisco: Jossey-Bass.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (1997). *Children, schools, and inequality.* Boulder, CO: Westview Press.

Freeman, D., & Freeman, Y. (1998). California reading: The pendulum swings. In K. S. Goodman (Ed.), *In defense of good teaching: What teachers need to know about the reading wars.* (pp. 73-85). York, ME: Stenhouse.

Gaddy, B. B., Hall, T. W., & Marzano, R. J. (1996). *School wars: Resolving our conflicts over religion and values.* San Francisco: Jossey-Bass.

General Accounting Office. (1993). *Educational achievement standards: NAGB's approach yields misleading interpretations.* GAO/PEMD-93-12. Washington, DC: Author.

Goodman, K. S. (1998). Who's afraid of whole language?: Politics, paradigms, pedagogy and the press. In K. S. Goodman (Ed.), *In defense of good teaching: What teachers need to know about the reading wars.* (pp. 3-37). York, ME: Stenhouse.

Guskey, T. R. & Oldham, B. R. (1997). Despite the best intentions: Inconsistencies among components in Kentucky's systemic reform. *Educational Policy*, 11, 426-442.

Hoffman, J., & McCarthy, S. (1995). The new basals: How are they different? *Reading Teacher, 49*, 72--75.

Hoffman, J. V., McCarthy, S. J., Elliott, B., Bayles, D., Price, D., Ferree, A., & Abbott, J. (1998). The literature-based basals in first-grade classrooms: Savior, Satan, or same-old, same-old? *Reading Research Quarterly, 33*, 168-197.

Holland, H. (1998). *Making change: Three educators join the battle for better schools.* Portsmouth, NH: Heinemann.

Jennings, N. E. (1996). *Interpreting policy in real classrooms: Case studies of state reform and teacher practice.* New York: Teachers College Press.

Johnston, P., Allington, R. L., Guice, S., & Brooks, G. W. (1998). Small change: A multi-level study of the implementation of literature-based instruction. *Peabody Journal of Education, 73*(3), 81-103.

Langer, J. A., & Allington, R. L. (1992). Curriculum research in writing and reading. In P. W. Jackson (Ed.), *Handbook of research on curriculum.* (pp. 687-725). New York: Macmillan.

Lindblom, C. E., & Woodhouse, E. J. (1993). *The policy-making process* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Loveless, T. (1998). The use and misuse of research in educational reform. In D. Ravitch (Ed.), *Brookings papers on educational policy, 1998.* Washington, DC: Brookings Institute.

Machan, D. (2000, January 24). Dr. Cram. *Forbes*, pp. 134-135.

McGill-Franzen, A. (2000). Policy and instruction: What is the relationship? In R. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (vol. 3, pp. 891-908). White Plains, NY: Longman.

McGill-Franzen, A. M. (1994). Is there accountability for learning and belief in children's potential? In E. H. Hiebert & B. M. Taylor (Eds.), *Getting reading right from the start: Effective early literacy interventions.* (pp. 13-35). Boston: Allyn-Bacon.

McLaughlin, M. W. (1991). The Rand Change Agent Study: Ten years later. In A. Odden (Ed.), *Education policy implementation* (pp. 143-155). Albany, NY: SUNY Press.

McQuillan, J. (1996). Whole language in California: A failed experiment? In P. H. Dreyer (Ed.), *Visions and realities in literacy: Yearbook of the 60th Claremont Reading Conference.* (pp. 17-26). Claremont, CA: Claremont Graduate School.

McQuillan, J. (1998). *The literacy crisis: False claims, real solutions.* Portsmouth, NH: Heinemann.

Miller, S. D. (1995). Teachers' responses to test-driven accountability: "If I change, will my scores drop?" *Reading Research and Instruction, 34*, 332-351.

National Center for Education Statistics. (1998). *The condition of education 1998.* Washington, DC: United States Government Printing Office.

National Education Commission on Time and Learning. (1994). *Prisoners of time.* Washington, DC: U.S. Government Printing Office.

Pace, G. (1992). Stories of teacher-initiated change from traditional to whole-language literacy instruction. *Elementary School Journal, 92*(4), 461-476.

Paterson, F. R. A. (1998). Mandating methodology: Promoting the use of phonics through state statute. In K. Goodman (Ed.), *In defense of good teaching: What teachers need to know about the "reading wars."* (pp. 107-125). York, ME: Stenhouse.

Paterson, F. R. A. (in press). The politics of phonics. *Journal of Curriculum Studies*.

Pearson, P. D. (1997, September). The politics of reading research and practice. *Council Chronicle*, p.8.

Pellegrino, J. W., Jones, L., & Mitchell, K. (1999). *Grading the nation's report card*. Committee on the Evaluation of National and State Assessments of Educational Progress, National Research Council. Washington, DC: National Academy Press. Available online at http://books.nap.edu/catalog/6296.html

Pressley, M., & Allington, R. L. (1999). What should educational research be the research of? *Issues in education: Contributions from educational psychology, 5*, 1-35.

Pressley, M., Wharton-McDonald, R., Allington, R. L., Block, C. C., Morrow, L., Tracey, D., Baker, K., Brooks, G., Cronin, J., Nelson, E., & Woo, D. (in press). The nature of effective first-grade literacy instruction. *Scientific Studies in Reading*.

Ravitch, D. (2000). Reading wars. In T. Loveless (Ed.), *The Curriculum Wars*. Washington, DC: Brookings Institution.

Ruddell, R. B. (1995). Those influential teachers: Meaning negotiators and motivation builders. *Reading Teacher, 48*, 454-463.

Scharer, P. L. (1992). Teachers in transitions: An exploration of changes in teachers and classrooms during the implementation of literature-based reading instruction. *Research in the Teaching of English, 2*, 408-443.

Shannon, P. (1996). Poverty, literacy, and politics: Living in the USA. *Journal of Literacy Research, 28*, 429-449.

Slavin, R. E. (1987). Making Chapter 1 make a difference. *Kappan, 69*, 110-119.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children: A report of the National Research Council*. Washington, DC: National Academy Press.

Spencer, L. M., & Spencer, S. M. (1993). *Competence at work: Models for superior performance*. New York: Wiley.

Spillane, J. P. (1998). A cognitive perspective on the role of the local educational agency in implementing instructional policy: Accounting for local variability. *Educational Administration Quarterly, 34*, 31-57.

Standerford, N. S. (1997). Reforming reading instruction on multiple levels: Interrelations and disconnections across the state, district, and classroom levels. *Educational Policy, 11*, 58-91.

Strickland, D. S., & Walmsley, S. A. (1993). *School book clubs and literacy development: A descriptive study*. M. R. Robinson Foundation.

Sweet, R. W. (1997, May/June). Don't read, don't tell: Clinton's phony war on illiteracy. *Policy Review*, pp. 38-42.

Taylor, D. (1998). *Beginning to read and the spin doctors of science: The political campaign to change America's mind about how children learn to read.* Urbana, IL: National Council of Teachers of English.

Timar, T. B., & Kirp, D. L. (1987). Educational reform and institutional competence. *Harvard Educational Review, 57*, 308-330.

Tyack, D., & Cuban, L. (1995). *Tinkering toward Utopia: A century of public school reform.* Cambridge, MA: Harvard University Press.

Urbanski, A. (1991, October 23). Real change is real hard: Lessons learned in Rochester. *Education Week*, p. 29.

Venezky, R. L. (1998). An alternate perspective on Success For All. In K. K. Wong (Ed.), *Advances in educational policy* (vol. 4, pp. 145-165). Greenwich, CT: JAI Press.

Vine, P. (1997, September 8). To market, to market: The school business sells kids short. *The Nation*, p. 265.

Walberg, H., & Greenberg, R. (1998, April 8). The Diogenes effect. *Education Week*, p. 60.

Wildavsky, A. (1979). *Speaking truth to power: The art and craft of policy analysis.* Boston: Little, Brown and Co.

Please help us assess the quality of our research report series by completing and returning the questionnaire below:

NAME OF REPORT: **EFFECTS OF READING POLICY ON CLASSROOM INSTRUCTION AND STUDENT ACHIEVEMENT**

1. Your position:

___ elementary school teacher	___ state ed. agency staff	___ policy maker
___ middle school teacher	___ professional developer	___ researcher
___ high school teacher	___ school administrator	___ education writer
___ college teacher/professor	___ district administrator	other_____

2. Clarity

	Very well				Not at all	
a. The concepts in this report were clearly expressed.	1	2	3	4	5	N/A
b. This report was well organized.	1	2	3	4	5	N/A

3. Utility

a. Reading this report gave me new information or insight into teaching or learning.	1	2	3	4	5	N/A
b. This report addresses a current and important problem in education,	1	2	3	4	5	N/A
. . . and offers a solution to the problem.	1	2	3	4	5	N/A
. . . and/or helps the reader understand the problem from a different perspective.	1	2	3	4	5	N/A
c. I found the ideas and solutions offered in this report to be feasible given current realities of policy and practice.	1	2	3	4	5	N/A

4. Scholarship

a. The conclusions drawn in the report are						
. . . adequately supported by the research presented.	1	2	3	4	5	N/A
. . . fully grounded in theory.	1	2	3	4	5	N/A

5. Any other comments or suggestions regarding this report or additional research needs in the area of English and language arts teaching and learning are greatly appreciated.

_____

_____

_____

_____

_____

Please return this form to:

Janet Angelis, Associate Director
Center on English Learning & Achievement
University at Albany, State University of New York
1400 Washington Ave., ED-B9
Albany, NY 12222 (518) 442-5023
jangelis@uamail.albany.edu